GISCOME – Genetics of Ischaemic Stroke Functional Outcome network: A protocol for an international multicentre genetic association study

GISCOME- Genetics of Ischemic Stroke Functional Outcome network: A protocol for an international multicentre genetic association study

Journal:	*European Stroke Journal*
Manuscript ID	ESO-16-0137.R1
Manuscript Type:	Protocol
Date Submitted by the Author:	27-Feb-2017
Complete List of Authors:	Maguire, Jane; University of Technology Sydney Faculty of Health Bevan, Steve; University of Lincoln, Joseph Banks Laboratories Stanne, Tara; Goteborgs universitet Sahlgrenska Akademin Lorentzen, Erik; Goteborgs universitet Sahlgrenska Akademin, Bioinformatics Core Facility Cadenas, Israel; Vall d'Hebron Institut de Recerca, Mediterranean building Neurovascular Research Lab Hankey, Graeme J.; University of Western Australia, School of Medicine and Pharmacology Jimenez-Conde, Jordi; Institut Hospital del Mar d'Investigació Mèdica, Department of Neurology Jood, Katarina; Goteborgs universitet Sahlgrenska Akademin Lee, Jin-Moo; Washington University in Saint Louis School of Medicine, Cerebrovascular Disease Lemmens, Robin; Katholieke Universiteit Leuven, Laboratory for Neurobiology (Vesalius Research Center) Levi, Christopher; University of Newcastle Priority Research Centre for Brain and Mental Health Research John Hunter Hospital Norrving, Bo Rannikmae, Kristiina ; University of Edinburgh Rost, Natalia; Harvard University Rosand, Jonathan; Massachusetts General Hospital, Neurology; Harvard Medical School Rothwell, Peter; University of Oxford Scott, Rodney; University of Newcastle Strbian, Daniel; University of Helsinki Sturm, Jonathan; University of Newcastle, Faculty of Health and Medicine Sudlow, Cathie; University of Edinburgh, Centre for Clinical Brain Sciences Traylor, Matthew; King's College London School of Medical Education Thijs, Vincent ; Austin Health Tatlisumak, Turgut; University of Gothenburg, Neurology Wieloch, Tadeusz ; Lunds Universitet, Faculty of Medicine Woo, Dan; University of Cincinnati Worrall, Bradford; University of Virginia Jern, Christina; Goteborgs universitet Sahlgrenska Akademin Lindgren, Arne; Skanes universitetssjukhus Lund Hematologiska kliniken, Department of Neurology and Rehabilitation Medicine
Keywords:	stroke, stroke recovery, genetic association, GWAMAS, functional outcome

Abstract:	Introduction: Genome-wide association (GWA) studies have identified several novel genetic loci associated with stroke risk, but how genetic factors influence stroke outcome is less studied. The Genetics of Ischemic Stroke Functional outcome (GISCOME) network aims at performing genetic studies of stroke outcome. We here describe the study protocol and methods basis of GISCOME. Methods: The GISCOME network has assembled patients from 12 ischemic stroke (IS) projects with genome-wide genotypic and outcome data from the International Stroke Genetics Consortium (ISGC) and the National Institute of Neurological Diseases Stroke Genetics Network (SiGN) initiatives. We have assessed the availability of baseline variables, outcome metrics and time-points for collection of outcome data. Results: We have collected 8831 IS cases with genotypic and outcome data. Modified Rankin score (mRS) was the outcome metric most readily available. We detected heterogeneity between cohorts for age and initial stroke severity (according to the NIH Stroke Scale), and will take this into account in analyses. We intend to conduct a first phase GWA outcome study on IS cases with data on initial stroke severity and mRS within 60-190 days. To date, we have assembled 5762 such cases and are currently seeking additional cases meeting these criteria for second phase analyses. Conclusion: GISCOME is a unique collection of IS cases with detailed genetic and outcome data providing an opportunity for discovery of genetic loci influencing functional outcome. GISCOME will serve as an exploratory study where the results as well as the methodological observations will provide a basis for future studies on functional outcome. GISCOME can also be used for candidate gene replication or assessing stroke outcome non-genetic association hypotheses.

GISCOME- Genetics of Ischemic Stroke Functional Outcome network: A protocol for an international multicentre genetic association study

Jane M. Maguire[1,2,3], Steve Bevan[4], Tara M. Stanne[5], Erik Lorenzen[5], Israel Fernandez-Cadenas[6,7], Graeme J Hankey[8], Jordi Jimenez-Conde[9,10], Katarina Jood[5], Jin-Moo Lee[11], Robin Lemmens[12,13,14], Christopher Levi[2,3,15], Bo Norrving[16,17], Kristiina Rannikmae[18], Natalia Rost[19], Jonathan Rosand[19,20,21], Peter M. Rothwell[22], Rodney Scott[1,2], Daniel Strbian[23], Jonathan Sturm[15], Cathie Sudlow[18], Matthew Traylor[24], Vincent Thijs[25], Turgut Tatlisumak[5,23], Tadeusz Wieloch[16], Daniel Woo[26], Bradford B. Worrall[27], Christina Jern[5]*, Arne Lindgren[16,17]*,

On behalf of the International Stroke Genetics Consortium and the NINDS-SiGN Consortium

*These authors contributed equally

Affiliations/Institutions:

1. Faculty of Health, University of Technology, Sydney, Australia
2. Hunter Medical Research Institute, University of Newcastle, Newcastle, Australia
3. Priority Research Centre for Stroke and Traumatic Brain Injury, University of Newcastle, Newcastle, Australia
4. School of Life Sciences, University of Lincoln, Lincoln, UK
5. Institute of Biomedicine, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden
6. Stroke Pharmacogenomics and Genetics, Fundació Docència I Recerca Mutuaterrassa, Mutua de Terrassa Hospital, Terrassa, Spain

7. Neurovascular Research Laboratory and Neurovascular Unit, Neurology and Medicine Departments – Universitat Autònoma de Barcelona, Vall d'Hebrón Hospital, Barcelona, Spain
8. School of Medicine and Pharmacology, The University of Western Australia, Perth, Australia
9. Department of Neurology. Institut Hospital del Mar d'Investigació Mèdica (IMIM), Barcelona, Spain
10. Universitat Pompeu Fabra, Barcelona, Spain
11. Department of Neurology, Washington University School of Medicine, St. Louis MO, USA
12. Department of Neurosciences, Experimental Neurology and Leuven Research Institute for Neuroscience and Disease (LIND), KU Leuven - University of Leuven, Leuven, Belgium
13. VIB, Vesalius Research Center, Laboratory of Neurobiology, Leuven, Belgium
14. Department of Neurology, University Hospitals Leuven, Leuven, Belgium
15. Faculty of Health and Medicine, University of Newcastle, Australia
16. Department of Clinical Sciences Lund, Neurology, Lund University, Lund, Sweden
17. Department of Neurology and Rehabilitation Medicine, Skane University Hospital, Lund, Sweden
18. Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK
19. Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston MA, USA
20. Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge MA, USA

21. Center for Human Genetic Research, Massachusetts General Hospital, Boston MA, USA
22. Stroke Prevention Research Unit, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK
23. Department of Neurology, Helsinki University Hospital, Helsinki, Finland
24. Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK
25. Department of Neurology, Austin Health, Heidelberg, Australia and Florey Institute for Neuroscience and Mental Health, University of Melbourne, Heidelberg, Australia
26. Department of Neurology and Rehabilitation University of Cincinnati College of Medicine, Cincinnati OH, USA
27. Departments of Neurology and Health Evaluation Sciences, University of Virginia, Charlottesville VA, USA

Corresponding author: Professor Jane Maguire

Jane.maguire@uts.edu.au

+61 02 9514 5971

Keywords: genetic association studies, stroke, functional outcome

Short title: A protocol for the GISCOME GWAS on ischemic stroke outcome

Word count: 4738 *Includes references, disclosures, tables and legends

Title page: 473

Abstract: 240

Main Body Text: 2612

Acknowledgements:

Sources of funding: Arne Lindgren is funded by: Region Skåne, Lund University, the Swedish Heart and Lung Foundation, the Freemasons Lodge of Instruction EOS Lund, and the Swedish Stroke Association. Christina Jern is funded by: the Swedish Research Council (K2014-64X-14605-12-5), the Swedish state and Region Västra Götaland (ALFGBG-429981), the Swedish Heart and Lung Foundation (20130315), and the Swedish Stroke Association. The Oxford Vascular Study has been funded by Wellcome Trust, Wolfson Foundation, UK Stroke Association, British Heart Foundation, Dunhill Medical Trust, and the NIHR Oxford Biomedical Research Centre. Peter Rothwell is in receipt of an NIHR Senior Investigator Award and a Wellcome Trust Senior Investigator Award.

Disclosures: None

References: 42

Tables: 2

Supplementary tables: 4

Supplementary figures: 1

Itemized list of tables and figures.

- Table 1. Characteristics of the 18 GISCOME cohorts.
- Table 2 The mRS distribution of ischaemic stroke patients in 60-190 day window for GISCOME cohorts intended for the primary analyses.

- Supplementary Table 1. Description and availability of data included in 18 GISCOME cohorts.
- Supplementary Table 2. Description of design methods used in 18 cohorts in GISCOME.
- Supplementary Table 3. Variables considered for inclusion in the primary GWA analyses for the 16 GISCOME cohorts planned for the primary analyses.
- Supplementary Table 4: Associations between age, sex and initial stroke severity (NIHSS) and functional outcome 3 months after stroke in the GISCOME data.
- Supplementary Figure. Simulated power calculations for genetic influence on functional outcome 3 months after stroke for a p-value < 5 x 10-8, based on the currently available data set.

Abstract:

Introduction: Genome-wide association (GWA) studies have identified several novel genetic loci associated with stroke risk, but how genetic factors influence stroke outcome is less studied. The **G**enetics of ***Is***chemic Stroke Functional out***come*** (GISCOME) network aims at performing genetic studies of stroke outcome. We here describe the study protocol and methods basis of GISCOME.

Methods: The GISCOME network has assembled patients from 12 ischemic stroke (IS) projects with genome-wide genotypic and outcome data from the International Stroke Genetics Consortium (ISGC) and the National Institute of Neurological Diseases Stroke Genetics Network (SiGN) initiatives. We have assessed the availability of baseline variables, outcome metrics and time-points for collection of outcome data.

Results: We have collected 8831 IS cases with genotypic and outcome data. Modified Rankin score (mRS) was the outcome metric most readily available. We detected heterogeneity between cohorts for age and initial stroke severity (according to the NIH Stroke Scale), and will take this into account in analyses. We intend to conduct a first phase GWA outcome study on IS cases with data on initial stroke severity and mRS within 60-190 days. To date, we have assembled 5762 such cases and are currently seeking additional cases meeting these criteria for second phase analyses.

Conclusion: GISCOME is a unique collection of IS cases with detailed genetic and outcome data providing an opportunity for discovery of genetic loci influencing functional outcome. GISCOME will serve as an exploratory study where the results as well as the methodological observations will provide a basis for future studies on functional outcome. GISCOME can also be used for candidate gene replication or assessing stroke outcome non-genetic association hypotheses.

Introduction

Globally, stroke is one of the principal causes of adult disability and the global burden of stroke is increasing (1, 2). After 1 year, up to 28% of stroke survivors are dependent on others for help with self-care and personal activities of everyday living (3). Even though last decades have shown significant reductions in stroke incidence in high-income countries, this has not been observed in low- or middle-income countries and with population aging and improved stroke survival, the absolute number of people who survive a stroke and experience varying levels of impairment continues to rise (1, 2). A deeper understanding of the biology of recovery after stroke is needed to identify new therapeutic targets for this affected group of patients.

Animal models demonstrate that following an acute ischaemic insult, the brain undergoes spontaneous recovery, repair, and remodelling (4). However, efforts to translate these findings to improve stroke outcomes in the clinical setting have been limited. Furthermore, the difficulty of predicting individual outcome poses a substantial challenge for ongoing post-stroke management strategies. Clinical parameters related directly to the acute event, such as age, stroke severity, etiologic stroke subtype, infarct size and location are predictors of outcomes (5-9), but predictive models based on these factors are imprecise (10-12). Other prognostic factors may include socioeconomic and social factors, post-stroke depression, and type and degree of treatment and rehabilitation (13) and there is a need for consensus on description of rehabilitation measures (14). Improvement of neurological function following the initial event is likely dependent on several of the above mentioned factors combined with environmental and genetic influences (15).

A genetic role in disease risk and susceptibility has been reported for many complex diseases including stroke (16-18), but the contribution of genetic factors to stroke outcomes is less

clear. There is substantial heritability reported for both intracerebral haemorrhage (ICH) and ischemic stroke (IS) (15, 19, 20). Preliminary evidence from individual candidate gene studies suggests that the functional outcome after stroke may also be determined by genetic factors in addition to clinical factors (21-25), however replication in larger cohorts is still outstanding. Genome wide association studies (GWASs) use designs that are hypothesis generating and have led to discovery of disease-associated loci across multiple phenotypes and subsequent new knowledge of genetic architecture of diseases (26). The **G**enetics of ***Is***chemic Stroke Functional out***come*** (GISCOME) effort therefore aims at detecting and describing genetic factors influencing IS outcomes, using data from already performed GWASs.

Here we describe the creation of the GISCOME network as the first international multi-centre collection of IS cases with data on outcomes, genome-wide genotypes, and salient baseline variables and the study protocol for future genetic analysis. We include a description of the process of selecting variables, outcome measures, and the potential future role of this collaboration network.

.

Methods

Twelve centers or joint projects have agreed to participate and provide data for analysis (Supplementary Table 1) and are already contributing to the International Stroke Genetics Consortium and the NINDS-SiGN Consortium efforts studying genetics of stroke risk. Some centers contributed more than one cohort of patients (eg. Barcelona) and some centers used multiple genotyping platforms (eg. Boston). This resulted in a total of 18 cohorts for which baseline characteristics, data availability and genotyping platform are outlined in Supplementary Table 1. The majority of the cohorts were hospital based with detailed phenotyping, including imaging. Supplementary Table 2 describes inclusion, recruitment period, and follow-up methods for each cohort. We have retrospectively collected phenotype data available for the 18 cohorts, selecting variables as outlined below.

Process of variable selection

The variables considered for inclusion in our study had already been collected in the individual cohorts by use of different study protocols. We conducted an initial survey across the cohorts to ascertain: 1) time-points when information on functional outcomes had been recorded; 2) what outcome measures had been utilised; and 3) all accessible baseline variables. We sought information on factors known or suspected to influence outcomes and these included: age, sex, living situation, stroke severity measured by National Institute of Health Stroke Scale (NIHSS) (26), ischemic stroke subtype, medical history/comorbid conditions, risk factors (including prior stroke or transient ischemic attack, coronary artery disease, atrial fibrillation, diabetes mellitus, hypertension, hyperlipidaemia, smoking, and alcohol use), pre-stroke physical functioning (measured with pre-stroke modified Rankin Score (mRS)), medications, and impairments and consequences of stroke such as cognitive

impairment and depression (27). We identified 71 variables and grouped them into: (1) Demographics, (2) Baseline characteristics, (3) Pre-stroke characteristics, (4) Risk factors, (5) Post-stroke treatments, (6) and Outcome measures. This provided us with a comprehensive overview of all variables available in at least one of the cohorts.

Next, we dropped the variables with unavailable data in more than one third of subjects. We selected modified Rankin Scale (mRS) score (27, 28) at 60-190 days as the most readily available functional outcome variable, after having observed that the majority of mRS values had been collected at 90 days +/- 2 weeks (81%) and that most of the remaining mRS observations were within the 60-190 day time span. The mRS values had already been scored by trained assessors at face-to-face or telephone follow-up for the majority of cohorts (for cohort specific details, please see Supplementary Table 2). The Lund Stroke Register and the Sahlgrenska Academy Study on Ischemic Stroke (SAHLSIS) phase 2 cohorts patients had been assessed with data from the 3 month follow up in the Swedish National Register Riksstroke. A validated algorithm for transforming answers on Riksstroke outcome questions into mRS grades was used even though this method prevented a differentiation between the mRS grades 0,1,2 (29). Baseline NIHSS was the selected measure for initial stroke severity. When multiple NIHSS scores were available we selected the score taken as close to 24 hours after stroke onset as possible (within 0-10 days).

Availability of ischemic stroke subtype classification data measured by Trial of ORG 10172 in Acute Stroke Treatment (TOAST) (30), Causative Classification System (CCS) (31, 32) or both, varied across the studies. Agreement between TOAST and CCS subtyping has been previously determined (31, 33) and there is significant genetic overlap between these two

methods (34) which suggests pooling cases with either classification may be beneficial in subsequent GISCOME studies.

Following this selection process, the phenotypic and genotypic data for each cohort were uploaded to central FTP secure servers located in Cambridge, UK, providing access to computational packages and file storage for this large-scale study. Further interrogation of the dataset led to a decision to remove additional variables not having a clear and homogeneous definition between cohorts or less than 50% availability. This included e.g. pre-stroke living and housing situation, ICH transformation after tPA, stroke to death interval, and recurrent stroke (Supplementary Table 3).

Meta-Analysis plan

Several assumptions were made by investigators and included consideration of the retrospective data from multiple cohorts and the subsequent limitations introduced by this. Thus, our planned analyses are considered as exploratory analyses to inform future prospective studies. We plan to analyse the primary outcome as mRS within the 60-190 days window, using first binary and then ordinal scales. The binary analyses will include both mRS 0-2 vs 3-6 and mRS 0-1 vs 2-6. For the analyses of mRS as an ordinal variable, ordinal logistic regression will be used. Simulated power calculations based on the currently available data are depicted in the Supplemental Figure. The ordinal model provides greater power. In this model with the available data set, the minimal odds ratios detectable at 80% power with a p-value $< 5 \times 10^{-8}$, are 1.15 for MAF 30%, 1.24 for MAF 10% and 1.35 for MAF 5%. Age, sex and initial NIHSS score are known to affect post-stroke outcome. To determine whether the expected associations were present in our data we performed regression analyses. As expected, all 3 of these variables were highly significant predictors of outcome in all 3 mRS

models described above (Supplementary Table 4). Therefore, in all analyses we will adjust for age and sex, and subsequently adjust for baseline NIHSS. We will adopt a standard GWAS significance threshold of 5×10^{-8} for all primary analyses. Because outcome results may depend on when evaluated after stroke onset we intend to do a sensitivity analysis for the majority group of our subjects with mRS outcome data available at 90 days +/- 2 weeks (81%). A separate secondary analysis including only subjects with baseline NIHSS available within 0-1 days is also planned.

Results

Characteristics of the GISCOME collection

We assembled a total of 8831 IS cases with phenotypic and genotype information in GISCOME. There were slightly fewer women (41.2%) than men, and cardiovascular risk factor frequencies were as expected in a stroke event group (Table 1). All cases included were of European ancestry and all cases were ≥18 years of age. Across all sites, stroke severity recorded at baseline were often mild strokes (NIHSS median 3; interquartile range (IQR), 1-7). Stroke severity was similar across the included cohorts, with the exception of three cohorts: VISP (median NIHSS 1, Interquartile range (IQR) 0-2); Val de Hebron-1 (median NIHSS 17, IQR 11-20); and Washington University (median NIHSS 8, IQR 4-12). It is of note that the median time of NIHSS scores for VISP were 70 days (IQR, 45-98). As only 0.3% of VISP fulfilled the NIHSS time window criteria of 0-10 days, this data set will not be included in the primary analysis. The distribution of TOAST subtypes was as follows: cardioembolic (CE) stroke 31.7%, large artery atherosclerotic (LAA) 17.9%, small vessel disease (SVD) 19.2%, other and undetermined 30.2% (Table 1). For CCS classification the distribution was: CE stroke 33.9%, LAA 16.4%, SVD 12.3% and other/underdetermined 37.4%. Loss to follow up ranged from 0-21% with the exception of the Massachusetts

General Hospital Genes Affecting Stroke Risk and Outcomes Study (MGH-GASROS) study (69% loss to follow-up) (Supplementary Table 2). The Edinburgh cohort subjects will not be included in the primary analysis because >90% lacked mRS outcome data within the 60-190 day window.

GISCOME subjects to be included in primary analyses

Given the considerations discussed above and the time windows selected for the primary GWA analysis (mRS day 60-190; NIHSS day 0-10), 5762 individuals from 16 cohorts are available with mRS, NIHSS and genotyping data for the primary analyses. Characteristics of this data set are summarized in Table 2. We intend to use the current dataset to conduct the first phase GWAS and then to expand to the second phase of this effort with data we expect to obtain from new cases from our existing studies and joint projects as well as from new contributing studies. A minimum set of variables required for phase 2 will include age, sex, stroke severity at 0-10 days and mRS at 60-190 days and available GWAS or DNA. Apart from the GWAS, we anticipate to specifically investigate known and putative genetic determinants of stroke outcome that include but are not limited to APOE and BDNF, both to validate these candidates and to demonstrate the viability of our cohort to replicate existing literature. We also plan to conduct the first GWAS based assessment for heritability of stroke outcome using a GWAS trait analysis approach using methods similar to those previously described regarding stroke risk (19). Insufficient sample size currently prevents the conduct of detailed subtype analyses at this stage, however we continue to seek additional cohorts to address this.

Discussion

This study protocol describes the GISCOME network which aims at conducting the first international multicentre large-scale GWAS on post-stroke outcomes. Within the GISCOME

cohorts, the most commonly used outcome metric was the mRS. Fortunately, this is one of the preferred functional outcome measures of choice in contemporary stroke trials (27). The mRS demonstrates strong test-retest and moderate inter-rater reliability which may possibly be enhanced by structured interviews and training (35-37). The clinical sensitivity or meaningful responsiveness to change in different outcome measures has been extensively studied (36). While mRS may not be the most sensitive scale to changes in functional activity, a one-point change in the scale is deemed to be clinically significant based on the range of activities captured by the scale (36).

Notably, the timing of outcome measures is equally important to the determination of outcome as the measure itself. By introducing time into consideration of outcome, two important derivative metrics emerge—the rate of change in outcome (rate of "recovery"), and maximal extent of outcome (extent of "recovery"). Rate of recovery refers to improvement per time unit. Extent of recovery refers to the functional ability, assessed by a metric such as mRS that captures the degree of functional ability. The biologic mechanism that underpins both of these is not well understood. However, because outcome metrics were not uniformly collected in several cohorts within GISCOME, we cannot currently study the rate of recovery. Improvement in functional outcome occurs most rapidly in the first days to weeks after ischemic brain injury; however, over the ensuing months, the degree of improvement plateaus (38). We chose to define mRS to encompass 60-190 days, but acknowledge that it is possible that some functional recovery may occur at earlier or later time points and this may not be accounted for in this investigation. The sensitivity analysis we propose will serve to determine how this may affect our results.

We selected age, sex, and initial stroke severity (as measured by baseline NIHSS within 0-10 days) as covariates in this analyses based on previous reports and our own observation that these variables influence functional outcome post-stroke. Study cohort also needs to be considered due to potential variability in outcomes due to differences in clinical practice specific to each stroke care system at the individual study sites. Other known determinants of post-stroke outcomes including pre-morbid status, acute stroke interventions (i.e., intravenous thrombolysis), neuroimaging characteristics of stroke (i.e., infarct size and location) will most likely not be included in this analysis due to lack of current data availability; however, ongoing studies within the ISGC such as MRI-GENIE (39) and TOTO (40) aim to provide additional information as to the role of specific stroke-related characteristics on genetics of functional outcomes in the future.

Our study has several strengths. We have assembled the largest sample of detailed stroke outcome phenotypic and genotypic data. The GISCOME network and proposed study will add to the understanding of genetic variants associated with neurological outcomes after the acute phase of ischemic stroke using individual level genetic data. Our retrospective design is largely pragmatic, taking advantage of existing datasets collected to examine stroke risk. The driving aim of GISCOME is to meta-analyse individual level data and identify novel genetic variants that influence the mechanistic pathways of functional outcomes post stroke. This parallels the efforts of other international consortia, several of which have extended the initial aim of identifying genetic risk factors associated with complex neurological disease to the investigation of genetic determinants of outcome e.g. Parkinson's Disease (41).

The retrospective design is a clear limitation, and introduces both selection and attrition bias since data included were previously collected under a variety of study protocols over a broad

time frame with notable loss to follow up in some cohorts. We thus had to derive our phenotypic data set from these heterogeneous sources. We selected the mRS at 60-190 days post index stroke as the primary outcome measure based on availability, and this metric is both acceptable and reliable in clinical stroke research (28, 35-37). However, while the mRS is widely acknowledged as the standard outcome measure in stroke clinical trials, we accept it is a relatively crude measure of functional recovery and the timing of mRS collection was not consistent across all contributing datasets. Even though data about mortality among the included subjects is available for the time of the primary outcome evaluation (i.e. as close to 90 days as possible), we do not have details about at the exact time point when deaths occurred. There was also heterogeneity between the individual cohorts regarding age and initial stroke severity and our total study sample has a bias towards milder strokes with median NIHSS of 3 which may hamper the detection of factors influencing the outcome in subjects with more severe stroke symptoms. We lacked data and/or clear definitions on several clinical variables known to influence outcome such as co-morbid depression, use of particular drugs e.g. selective serotonin reuptake inhibitors or anticonvulsants, and measures of social support (Supplementary Table 3). We also lack data on the volume of infarct, but as infarct volume is known to correlate with NIHSS we will be able to partly account for this. Finally, all cases were of European ancestry and do not represent a global stroke population. Therefore, specific genetic factors influencing outcome after stroke in subjects with other ethnic backgrounds will not be detected. We aim to address this in future efforts. In phase 2 we will seek and invite sites that are derived from more diverse ethno-geographic groups. In the future, an expansion of the number of study subjects is also needed to improve the power of detecting genetic variants related to ischemic stroke outcome. Despite these limitations, a major strength of our planned analysis is the detailed description of the methods used and

careful selection of a much needed repository for novel investigation into genetic determinants of stroke outcome.

Conclusion

The GISCOME study protocol describes an exploratory effort providing an excellent opportunity to detect genetic influence on stroke outcomes and to inform future studies within this important field of stroke research. The GISCOME sample size will increase through identification of additional sites and recruitment of cases within existing studies. We anticipate that this will increase our capability to explore other avenues of inquiry, for example, variants of smaller effect sizes.

We also strongly advocate for future prospective cohorts to utilize measures of functional capacity, quality of life, and neuropsychological function. We therefore urge the stroke community to characterize stroke cases using standardised definitions (42) and follow up stroke patients in their acute and rehabilitation phases with consistent documentation of functional ability. Co-operation within e.g. the International Stroke Genetics Consortium is an effective method to increase the availability of studies for this type of research. These efforts will provide a stable platform for identifying genetic variants that are associated with functional outcome.

References:

1. Feigin VL, Forouzanfar MH, Krishnamurthi R, Mensah GA, Connor M, Bennett DA, et al. Global and regional burden of stroke during 1990-2010: findings from the Global Burden of Disease Study 2010. Lancet. 2014;383(9913):245-54.
2. Feigin VL, Krishnamurthi RV, Parmar P, Norrving B, Mensah GA, Bennett DA, et al. Update on the Global Burden of Ischemic and Hemorrhagic Stroke in 1990-2013: The GBD 2013 Study. Neuroepidemiology. 2015;45(3):161-76.
3. Ullberg T, Zia E, Petersson J, Norrving B. Changes in functional outcome over the first year after stroke: an observational study from the Swedish stroke register. Stroke. 2015;46(2):389-94.
4. Overman JJ, Carmichael ST. Plasticity in the injured brain: more than molecules matter. The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry. 2014;20(1):15-28.

5. Coupar F, Pollock A, Rowe P, Weir C, Langhorne P. Predictors of upper limb recovery after stroke: a systematic review and meta-analysis. Clinical rehabilitation. 2012;26(4):291-313.
6. Carter AM, Catto AJ, Mansfield MW, Bamford JM, Grant PJ. Predictive variables for mortality after acute ischemic stroke. Stroke. 2007;38(6):1873-80.
7. Muscari A, Puddu GM, Santoro N, Zoli M. A simple scoring system for outcome prediction of ischemic stroke. Acta neurologica Scandinavica. 2011;124(5):334-42.
8. Cramer SC. Repairing the human brain after stroke: I. Mechanisms of spontaneous recovery. Ann Neurol. 2008;63(3):272-87.
9. Abdul-Rahim AH, Quinn TJ, Alder S, Clark AB, Musgrave SD, Langhorne P, et al. Derivation and Validation of a Novel Prognostic Scale (Modified-Stroke Subtype, Oxfordshire Community Stroke Project Classification, Age, and Prestroke Modified Rankin) to Predict Early Mortality in Acute Stroke. Stroke. 2016;47(1):74-9.
10. Counsell C, Dennis M, McDowall M, Warlow C. Predicting outcome after acute and subacute stroke: development and validation of new prognostic models. Stroke. 2002;33(4):1041-7.
11. Strbian D, Meretoja A, Ahlhelm FJ, Pitkaniemi J, Lyrer P, Kaste M, et al. Predicting outcome of IV thrombolysis-treated ischemic stroke patients: the DRAGON score. Neurology. 2012;78(6):427-32.
12. Saposnik G. The art of estimating outcomes and treating patients with stroke in the 21st century. Stroke. 2014;45(6):1603-5.
13. Hankey GJ. Stroke. Lancet. 2016;Published Online September 13, 2016 doi: 10.1016/S0140-6736(16)30962-X. [Epub ahead of print].
14. Bernhardt J, Borschmann K, Boyd L, Thomas Carmichael S, Corbett D, Cramer SC, et al. Moving rehabilitation research forward: Developing consensus statements for rehabilitation and recovery research. International journal of stroke : official journal of the International Stroke Society. 2016;11(4):454-8.
15. Lindgren A, Maguire J. Stroke Recovery Genetics. Stroke. 2016;47(9):2427-34.
16. Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. Lancet Neurol. 2012;11(11):951-62.
17. McCarthy MI, Zeggini E. Genome-wide association studies in type 2 diabetes. Current diabetes reports. 2009;9(2):164-71.
18. Holliday EG, Maguire JM, Evans TJ, Koblar SA, Jannes J, Sturm JW, et al. Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. Nat Genet. 2012;44(10):1147-51.
19. Bevan S, Traylor M, Adib-Samii P, Malik R, Paul NL, Jackson C, et al. Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. Stroke. 2012;43(12):3161-7.
20. Devan WJ, Falcone GJ, Anderson CD, Jagiella JM, Schmidt H, Hansen BM, et al. Heritability estimates identify a substantial genetic contribution to risk and outcome of intracerebral hemorrhage. Stroke. 2013;44(6):1578-83.
21. Cramer SC, Procaccio V. Correlation between genetic polymorphisms and stroke recovery: analysis of the GAIN Americas and GAIN International Studies. Eur J Neurol. 2012;19(5):718-24.
22. Maguire J, Thakkinstian A, Levi C, Lincz L, Bisset L, Sturm J, et al. Impact of COX-2 rs5275 and rs20417 and GPIIIa rs5918 polymorphisms on 90-day ischemic stroke functional

outcome: a novel finding. Journal of stroke and cerebrovascular diseases : the official journal of National Stroke Association. 2011;20(2):134-44.
23. Aberg ND, Olsson S, Aberg D, Jood K, Stanne TM, Nilsson M, et al. Genetic variation at the IGF1 locus shows association with post-stroke outcome and to circulating IGF1. European journal of endocrinology / European Federation of Endocrine Societies. 2013;169(6):759-65.
24. Hoy A, Leininger-Muller B, Poirier O, Siest G, Gautier M, Elbaz A, et al. Myeloperoxidase polymorphisms in brain infarction. Association with infarct size and functional outcome. Atherosclerosis. 2003;167(2):223-30.
25. Liepert J, Heller A, Behnisch G, Schoenfeld A. Catechol-O-methyltransferase polymorphism influences outcome after ischemic stroke: a prospective double-blind study. Neurorehabilitation and neural repair. 2013;27(6):491-6.
26. Adams HP, Jr., Davis PH, Leira EC, Chang KC, Bendixen BH, Clarke WR, et al. Baseline NIH Stroke Scale score strongly predicts outcome after stroke: A report of the Trial of Org 10172 in Acute Stroke Treatment (TOAST). Neurology. 1999;53(1):126-31.
27. Katzan IL, Spertus J, Bettger JP, Bravata DM, Reeves MJ, Smith EE, et al. Risk adjustment of ischemic stroke outcomes for comparing hospital performance: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2014;45(3):918-44.
28. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. Stroke. 1988;19(5):604-7.
29. Eriksson M, Appelros P, Norrving B, Terent A, Stegmayr B. Assessment of functional outcome in a national quality register for acute stroke: can simple self-reported items be transformed into the modified Rankin Scale? Stroke. 2007;38(4):1384-6.
30. Adams HP, Jr., Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. Stroke. 1993;24(1):35-41.
31. Ay H, Furie KL, Singhal A, Smith WS, Sorensen AG, Koroshetz WJ. An evidence-based causative classification system for acute ischemic stroke. Ann Neurol. 2005;58(5):688-97.
32. Ay H, Benner T, Arsava EM, Furie KL, Singhal AB, Jensen MB, et al. A computerized algorithm for etiologic classification of ischemic stroke: the Causative Classification of Stroke System. Stroke. 2007;38(11):2979-84.
33. McArdle PF, Kittner SJ, Ay H, Brown RD, Jr., Meschia JF, Rundek T, et al. Agreement between TOAST and CCS ischemic stroke classification: the NINDS SiGN study. Neurology. 2014;83(18):1653-60.
34. (NINDS), (ISGC). Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. Lancet Neurol. 2016;15:174-84.
35. Quinn TJ, Dawson J, Walters MR, Lees KR. Reliability of the modified Rankin Scale: a systematic review. Stroke. 2009;40(10):3393-5.
36. Harrison JK, McArthur KS, Quinn TJ. Assessment scales in stroke: clinimetric and clinical considerations. Clinical interventions in aging. 2013;8:201-11.
37. Banks JL, Marotta CA. Outcomes validity and reliability of the modified Rankin scale: implications for stroke clinical trials: a literature review and synthesis. Stroke. 2007;38(3):1091-6.

38. Kwakkel G, Kollen BJ. Predicting activities after stroke: what is clinically relevant? International journal of stroke : official journal of the International Stroke Society. 2013;8(1):25-32.
39. Rost NS. MRI-GENIE study 2015 [cited 2016 Dec 11]. Available from: . http://360bio/grants/88888898768/mri-genetics-interface-exploration-mri-genie-study/.
40. Holliday E, Maguire J, Thijs V, Koblar S, Sturm J, Lincz L. Helping stroke physicians choose who to thrombolyse - the "Targeting Optimal Thrombolysis Outcomes" (TOTO) study 2015 [cited 2016 Dec 11]. Available from: . https://researchdataandsorgau/helping-stroke-physicians-toto-study/519093.
41. Chung SJ, Armasu SM, Biernacka JM, Anderson KJ, Lesnick TG, Rider DN, et al. Genomic determinants of motor and cognitive outcomes in Parkinson's disease. Parkinsonism & related disorders. 2012;18(7):881-6.
42. Majersik JJ, Cole JW, Golledge J, Rost NS, Chan YFY, Gurol ME, et al. Recommendations from the international stroke genetics consortium, part 1: Standardized phenotypic data collection. Stroke. 2015;46(1):279-84.

Table 1 Characteristics of the 18 GISCOME cohorts

Characteristics	Ischaemic stroke cases (number = 8831)	Missing data number (%)
mRS available (60-190d), number (%)	7416 (84.0)	1415 (16.0)
mRS taken at day, median (IQR)	90 (81-90)	
Sex, female number (%)	3658 (41.4)	0 (0)
Age, years mean (SD)	68.4 (13.5)	0 (0)
NIHSS available (0-10d), number (%)**	6820 (77.2)	2011 (22.8)
NIHSS, median (IQR)	3 (1-7)	
NIHSS taken at day, median (IQR)	0 (0-1)	
Rehabilitation measures registered, number/available (%)	1638/3387 (48.4)	5444 (61.6)
TOAST Stroke subtypes, number/available (%)		
Cases with TOAST data	6437 (72.9)	2394 (27.1)
Large artery atherosclerosis	1155/6437 (17.9)	
Cardioembolic	2038/6437 (31.7)	
Small vessel disease	1235/6437 (19.2)	
Other/undetermined	2009/6437 (31.2)	
CCS Stroke subtypes, number/available (%)		
Cases with CCS data	4694 (53.2)	4137 (46.8)
Large artery atherosclerosis	770/4694 (16.4)	
Cardioembolic	1593/4694 (33.9)	
Small vessel disease	576/4694 (12.3)	
Other/undetermined	1755/4694 (37.4)	
Cardiovascular risk factors, number/available (%)*		
Hypertension	5891/8787 (67.0)	44 (0.5)
Hypercholesterolemia	4715/8530 (55.3)	301 (3.4)
Diabetes mellitus	1940/8622 (22.5)	209 (2.4)
Atrial fibrillation	1746/8799 (19.8)	32 (0.4)
Ischaemic Heart Disease	1589/7474 (21.3)	1357 (15.4)
Current smoker	2007/8683 (23.1)	148 (1.7)
Pharmacological intervention,		
Cases treated with Alteplase, number/available (%)	689/4886 (14.1)	3945 (44.7)
Premorbid impaired functional status, number (%)	772/6867 (77.8)	1964 (22.2)
Pre-stroke living situation, number/available (%)		6095 (69)
Alone	897/2736 (32.8)	
Divorced	64/2736 (2.3)	
Widowed	17/2736 (0.6)	
With someone	1758/2736 (64.3)	
First or recurrent stroke, number/available (%)		477 (5.4)
First	6797/8354 (81.4)	
Recurrent	1557/8354 (18.6)	
Pre-stroke housing, number/available (%)		6430 (72.8)
Assisted living	5/2401 (0.2)	
Institution	55/2401 (2.3)	
Nursing home	13/2401 (0.5)	
Own house/flat	2318/2401 (96.5)	
Other	10/2401 (0.4)	

* Availability across cohorts. Numbers vary per cohort.

** Only n=5/1723 (0.3%) of individuals in VISP fulfilled the NIHSS time window criteria of 0-10d.

mRS indicates modified Rankin Scale; IQR, interquartile range; SD, standard deviation; NIHSS, NIH stroke scale; TOAST, Trial of Org 10172 in Acute Stroke Treatment stroke sub classification; CCS, Causative Classification System

Table 2. The mRS distribution of ischaemic stroke patients in 60-190 day window for GISCOME cohorts intended for the primary analyses.

Characteristics	16 cohorts Ischaemic stroke cases (number = 5762)	14 cohorts* Ischemic stroke cases (number = 4421)
Sex, females number (%)	2472 (42.9)	1894 (42.8)
Age, years mean (SD)	68.6 (14.0)	68.7 (13.9)
mRS dichotomized 0-2 vs 3-6		
Poor outcome, number (%)	2131 (37.0)	N/A
mRS dichotomized 0-1 vs 2-6		
Poor outcome, number (%)	N/A	2567 (58.1)
mRS, ordinal scale		
0	718 (12.5)	718 (16.2)
1	1953 (33.9) **	1136 (25.7)
2	960 (16.7)	960 (21.7)
3	847 (14.7)	628 (14.2)
4	605 (10.5)	479 (10.8)
5	215 (3.7)	138 (3.1)
6	464 (8.1)	362 (8.2)
NIHSS (0-10 days), median (IQR)	4 (2-9)	4 (2-9)
NIHSS taken at day, median (IQR)	0 (0-1)	0 (0-1)

*LSR and SAHLSIS Phase 2 not included in this distribution because these cohorts used a collapsed score for mRS 0-2. **Including LSR and SAHLSIS phase 2, where collapsed mRS scores of 0-2 (number=519 and number=298 subjects, respectively) are assigned as having mRS 1. SD indicates standard deviation; mRS, modified Rankin Scale, LSR, Lund Stroke Register; SAHLSIS, Sahlgrenska Academy Study on Ischemic Stroke; NIHSS, NIH stroke scale; N/A, not applicable; IQR, interquartile range.

Supplementary Table 1. Description and availability of data included in the 18 GISCOME cohorts.

Center/ Joint project	Cohort	N	Age (mean)	Female (%)	NIHSS at baseline (median [IQR])	Genotype platform	mRS at 60-190 days (%)
Australia	VIS	580	71.3	45.2	4 (2-8)	Illumina 610Q	86.7
Barcelona	Hospital del Mar[1]	924	75.2	47	5 (3-12)	Illumina 5M	98.1
	Val de Hebron-1	170	73.0	52.4	17 (11-20)	Illumina Omni1-Quad	77.1
	Val de Hebron-2	105	68.2	25.7	3 (1.5-8)	Illumina Omni2.5- Quad	100
Boston	MGH-GASROS	158	66.9	38.6	3 (1-9)	Illumina 610Q	68.4
	MGH-GASROS	169	68.7	46.7	3 (2-9)	Affymetrix 6.0	29.0
	MGH-GASROS[1]	527	64.5	35.9	3 (1-7)	Illumina 5M	48.8
Cincinnati	GCNKSS[1]	372	69.4	44.9	4 (2-8)	Illumina 5M	95.2
Edinburgh	The Edinburgh Stroke Study*	483	70.9	44.7	4 (1-7)	Illumina 660Q	9.9
Gothenburg	SAHLSIS	261	59.3	41.8	3 (2-7)	Illumina 750K	88.1
	SAHLSIS[1]	797	54.6	33.4	3 (1-8)	Illumina 5M	88.8
Helsinki	Helsinki-2000 Study	351	63.9	38.5	5 (2-10)	Illumina CoreExome	100
Leuven	LSGS[1]	469	67.5	41.4	4 (2-8)	Illumina 5M	97.7
Lund	LSR	528	74.3	47.7	3 (2-7)	Illumina 750K	92.6
	LSR[1]	574	71.5	44.9	4 (2-8)	Illumina 5M	83.4
Oxford	Oxford Stroke Study	548	74.0	50.5	2 (0-4)	Illumina 660Q	98.9
Virginia	VISP*	1723	68.0	35.0	1 (0-2)	Illumina 1M	93.2
Washington	WASH-U[1]	92	67.2	43.5	8 (4-12)	Illumina 5M	100

*not included in the primary GISCOME analyses. [1]Genotyped in the Stroke Genetic Network (SiGN) study. N indicates number; NIHSS, National Institute Health Stroke Scale; IQR, Interquartile Range ; mRS, modified Rankin score ; VIS, indicates Vascular Ischemia Study; MGH- GASROS, Massachusetts General Hospital Genes Affecting Stroke Risk and Outcomes Study; GCNKSS, Greater Cincinnati/ Northern Kentucky Stroke Study; SAHLSIS, Sahlgrenska Academy Study on Ischemic Stroke; LSGS, Leuven Stroke Genetics Study; LSR, Lund Stroke Register; VISP, Vitamin Intervention for Stroke Prevention study; WASH-U, Washington University Stroke Study.

Supplementary Table 2. Description of design methods used in the 18 GISCOME cohorts.

Center/ Joint project	Cohort(s)	Age Range (years)	Study Design	Recruitment period	Follow-up mRS	Estimated loss to follow up
Australia	VIS	>18	Hospital based	2003-2006	Telephone	15%
Barcelona	Hospital del Mar Val de Hebron-1 and -2	All All	Hospital based	2011-ongoing	Telephone	0%
Boston	MGH- GASROS	≥18	Hospital based	2001-2011	Telephone	69%
Cincinnati	GCNKSS	All	Population based	1993-2010	Face-to-face	15%
Edinburgh	The Edinburgh Stroke Study	>18	Hospital based	2002-2005	Postal questionnaire	5%
Gothenburg	SAHLSIS, phase 1	18-69	Hospital based	1998-2003	Face-to-face	5%
Gothenburg	SAHLSIS, phase 2	18-69	Hospital based	2004-2011	Riksstroke register	21%
Helsinki	Helsinki-2000 Study	>18	Hospital based	2011-ongoing	Face-to-face	0%
Leuven	LSGS	≥18	Hospital based	2008	Face-to-face	2.2%
Lund	LSR	≥18	Hospital based	2001-ongoing	Riksstroke register	12%
Oxford	Oxford Stroke Study	All	Community based	2002-2004	Face-to-face	0%
Virginia	VISP	>35	Multi centre, double blind, RCT	1996-2003	Face-to-face	6%
Washington	WASH-U	>18	Hospital based*	2008-2013	Telephone	0%

When a center used the same methods for several of their cohorts, these are collapsed to one row in this table. *Only patients with NIHSS at 24hrs = 4-20 and Baseline mRS < 2 included. VIS indicates Vascular Ischemia Study; MGH- GASROS, Massachusetts General Hospital Genes Affecting Stroke Risk and Outcomes Study; GCNKSS, Greater Cincinnati/ Northern Kentucky Stroke Study; SAHLSIS, Sahlgrenska Academy Study on Ischemic Stroke; LSGS, Leuven Stroke Genetics Study; LSR, Lund Stroke Register; VISP, Vitamin Intervention for Stroke prevention; WASH-U, Washington University Stroke Study, NIHSS, NIH stroke scale.

Supplementary Table 3. Variables considered for inclusion in the primary GWAS analyses for the 16 GISCOME cohorts intended for the primary analyses.

Variable group	Variable	Description	Available*
Demographics	Age	numerical (years)	yes
	Gender	Female/male	yes
Pre-stroke characteristics	Premorbid functional status	Normal/impaired	yes
	Serious comorbidity pre-stroke	IHD or other/no	no
	Living situation pre-stroke	Alone/with someone	no
	Housing pre-stroke	Own house or flat/institution	no
Baseline	First/recurrent stroke	First/recurrent	yes
	TOAST	LAA, CE, SVD, UNK	yes
	CCS	LAA, CE, SVD, UNK	yes
	NIHSS	numerical	yes
	tPA therapy in acute phase	yes/no	yes
	ICH transformation after tPA treatment	yes/no	no
	Recanalisation approximately 1hour after tPA	yes/no	no
Risk factors	Hypertension	yes/no	yes
	Diabetes mellitus	yes/no	yes
	Current smoking	yes/no/ex	yes
	Atrial fibrillation	yes/no	yes
	Ischemic heart disease	yes/no	yes
	Hypercholesterolemia	yes/no	yes
Treatment post stroke	Discharge to	Rehabilitation/other hospital/nursing home/ other institution/home/ dead	yes
	Rehabilitation treatment	yes/no	yes
	SSRI or similar treatment	yes/no	no
Outcome	Interval between stroke and death	numerical (days)	no
	Recurrent stroke	yes/no	no
	Depression	yes/no	no
	NIHSS at various time points	numerical	yes at baseline
	mRS at various time points	numerical	yes at 3 months (60-190 days)
	Glasgow Outcome Scale at various time points	numerical	no
	Barthel Index at various time points	numerical	no

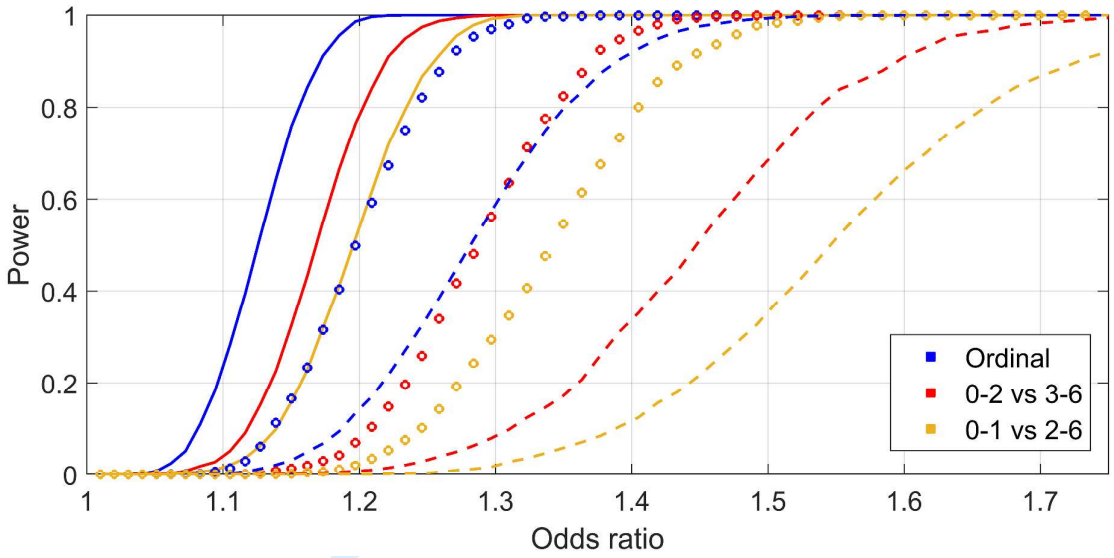
*Variable having clear and homogeneous definitions between cohorts and available in at least 50% of subjects. GWAS indicates genome wide association study; IHD, ischaemic heart disease; TOAST, Trial of Org 10172 in Acute Stroke Treatment stroke sub classification; CCS, Causative Classification System; LAA, Large artery atherosclerosis; CE, Cardioembolic; SVD, Small vessel disease; UNK, Other/undetermined; NIHSS, NIH stroke scale; tPA, tissue plasminogen activator; ICH, intracerebral haemorrhage; SSRI, Selective serotonin reuptake inhibitor; mRS, modified Rankin Scale.

Supplementary Table 4: Associations between age, sex and initial stroke severity (NIHSS) and functional outcome (mRS) 3 months after stroke in the GISCOME data.

	mRS 0-2 vs 3-6 OR (95% CI), *P*	mRS 0-1 vs 2-6 OR (95% CI), *P*	Ordinal (0-6) OR (95% CI), *P*
Age	1.06 (1.05-1.06), $1.6x10^{-122}$	1.03 (1.03-1.03), $2.5x10^{-48}$	1.04 (1.04-1.04), $1.4x10^{-119}$
Sex (female)	1.72 (1.54-1.91), $1.1x10^{-22}$	1.55 (1.40-1.73), $1.6x10^{-16}$	1.68 (1.53-1.85), $9.7x10^{-26}$
NIHSS (0-10d)	1.16 (1.15-1.17), $2.7x10^{-184}$	1.14 (1.13-1.15), $2.0x10^{-119}$	1.19 (1.17-1.20), $1.4x10^{-297}$

mRS, modified Rankin Scale; OR, odds ratio; CI, confidence intervals. OR is for poor outcome per year for age, for females, and per 1 score for NIHSS.

Supplementary Figure. Simulated power calculations for genetic influence on functional outcome 3 months after stroke for a p-value < 5 x 10-8, based on the currently available data set.



Blue indicates ordinal regression for 16 GISCOME cohorts (n= 5762); red indicates binary regression (mRS 0-2 vs 3-6) for 16 GISCOME cohorts (n= 5762); yellow indicates binary regression (mRS 0-1 vs 2-6) logistic regression for 14 GISCOME cohorts (n= 4421; LSR and SAHLSIS Phase 2 are not included in this distribution because these cohorts used a collapsed mRS score 0-2). Lines indicate simulated power for minor allele frequency (MAF) 0.3, open circles for MAF 0.1, and dashed line for MAF 0.05. The mRS was scored within a 60-190 day window.

DECLARATIONS

Conflicting interests: All Authors declare that there is no conflict of interest.

Disclosures: Robin Lemmens is a senior clinical investigator of FWO Flanders. All other authors had no disclosures.

Sources of funding: Christina Jern: the Swedish Research Council (K2014-64X-14605-12-5), the Swedish state and Region Västra Götaland (ALFGBG-429981), the Swedish Heart and Lung Foundation (20130315), and the Swedish Stroke Association. Arne Lindgren: Region Skåne, Lund University, the Swedish Heart and Lung Foundation, the Freemasons Lodge of Instruction EOS Lund, and the Swedish Stroke Association. Helsinki-2000 Study is partly funded by grants from the Helsinki University Central Hospital and the Sigrid Juselius Foundation. Turgut Tatlisumak: Helsinki-2000 ethics permit: 27-OCT-2010, 266, 287/13/03/01/2010. Peter Rothwell: The Oxford Vascular Study has been funded by Wellcome Trust, Wolfson Foundation, UK Stroke Association, British Heart Foundation, Dunhill Medical Trust, and the NIHR Oxford Biomedical Research Centre. Professor Rothwell is in receipt of an NIHR Senior Investigator Award and a Wellcome Trust Senior Investigator Award. Robin Lemmens is a senior clinical investigator of FWO Flanders. Daniel Strbian was supported with a grant from the Finnish Medical Foundation and Helsinki University Hospital governmental subsidiary funds for clinical research.

Informed consent: Written informed consent was obtained from the patient(s) for their anonymized information to be published in this article.

Ethical approval: All studies included in this manuscript were granted approval by their local ethics committees.

Guarantor: Jane Maguire. All co-authors, confirmed by email “I consent that Dr Jane Maguire submit the article for publication and I give her the right to sign the Contributor Agreement on my behalf”.

Contributorship: JM, AL, CJ researched literature and conceived the study. JM, JML,RL,CS, RS, BR, CL,VT, EL, BN, CJ, was involved in protocol development, JR, CJ,VT,RL, TT,TW, CL, JM, TS, VT gaining ethical approval, JM, TS, EL, VT, SB, BBW, JR patient recruitment and data analysis. JM, AL, CJ wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript

Acknowledgements: We would like to thank Prof John Attia and A/Prof Elizabeth Holliday for his/her assistance and guidance in this research.